

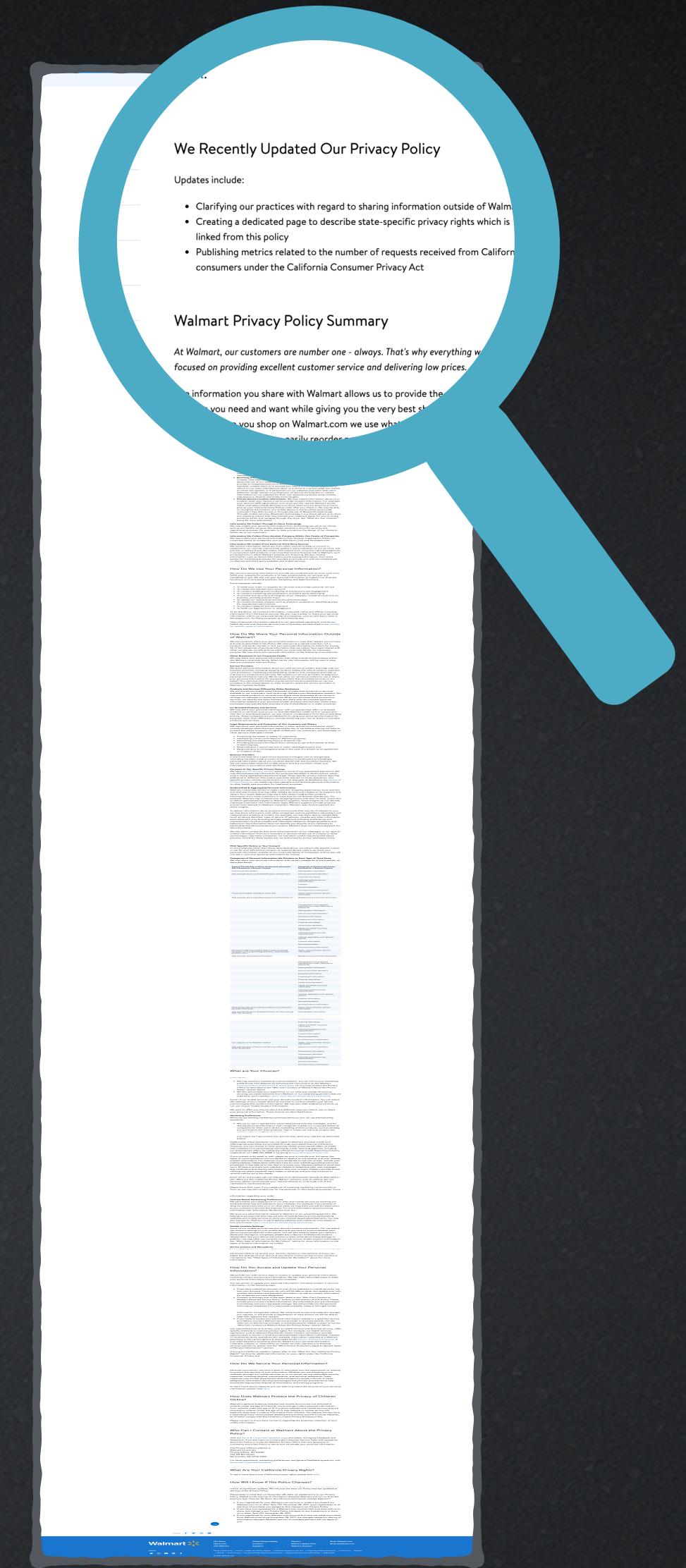
# Detecting Textual Saliency in Privacy Policy

---



Phase-1 Presentation

# Recap: A Close Look At Privacy Policy



## Details about

- ? How and why do they collect our information?
- ? Are they sharing this data with third parties?
- ? How long is my data stored?
- ? Can I access my data?
- ? How secure and protected is my data?

# Goal

Provide key information to end-users using Machine Learning and Natural Language Processing



## INFORM Module

- ★ Visualization
- ★ Annotating policies



## QUERY Module

- ★ Enable users to ask policy-related

Interactive

## Web Framework



# Stratified Splitting

- Goal: Pursue an unbiased measuring approach
- Split the dataset into 70% train, 15% validation, 15% test

Traditional Split

SD: 25.43

	(0,)	(1,)	(10,)	(11,)	(2,)	(3,)	(4,)	(5,)	(6,)	(7,)	(8,)	(9,)
Train	106	257	157	450	23	1062	233	576	138	460	216	850
Val	98	266	163	373	18	1017	322	592	135	410	228	840
Test	135	284	177	476	23	1115	238	513	116	438	270	728

Iterative Stratification

SD: 1.83

	(0,)	(1,)	(10,)	(11,)	(2,)	(3,)	(4,)	(5,)	(6,)	(7,)	(8,)	(9,)
Train	109	262	161	442	22	1063	247	569	134	449	226	830
Val	116	266	163	443	32	1064	247	569	135	448	228	830
Test	102	261	158	443	14	1064	247	569	135	452	224	830

# Split Statistics (Union)

Number of unique segments in total: 3776

	TRAIN SET			DEV SET		
	Counts	Percentage	Counts	Percentage	Counts	Percentage
Data Retention	109	2.41%	24	2.47%	23	2.37%
Data Security	262	5.8%	56	5.77%	57	5.88%
Do Not Track	22	0.49%	5	0.52%	5	0.52%
First Party Collection/Use	1063	23.55%	228	23.51%	228	23.53%
International and Specific Audiences	247	5.47%	53	5.46%	53	5.47%
Introductory/Generic	569	12.61%	122	12.58%	122	12.59%
Policy Change	134	2.97%	29	2.99%	29	2.99%
Practice not covered	449	9.95%	96	9.9%	97	10.01%
Privacy contact information	226	5.01%	49	5.05%	48	4.95%
Third Party Sharing/Collection	830	18.39%	178	18.35%	178	18.37%
User Access, Edit and Deletion	161	3.57%	35	3.61%	34	3.51%
User Choice/Control	442	9.79%	95	9.79%	95	9.8%
	2645		563		568	

# Split Statistics (Majority)

Number of unique segments in total: 3471

TRAIN SET    DEV SET    TEST SET

	Counts	Percentage	Counts	Percentage	Counts	Percentage
Data Retention	55	1.91%	11	1.79%	12	1.94%
Data Security	147	5.1%	32	5.19%	31	5.0%
Do Not Track	22	0.76%	4	0.65%	5	0.81%
First Party Collection/Use	845	29.3%	181	29.38%	181	29.19%
International and Specific Audiences	211	7.32%	45	7.31%	45	7.26%
Introductory/Generic	273	9.47%	58	9.42%	59	9.52%
Policy Change	83	2.88%	18	2.92%	18	2.9%
Practice not covered	90	3.12%	19	3.08%	20	3.23%
Privacy contact information	142	4.92%	31	5.03%	30	4.84%
Third Party Sharing/Collection	661	22.92%	142	23.05%	142	22.9%
User Access, Edit and Deletion	104	3.61%	22	3.57%	23	3.71%
User Choice/Control	251	8.7%	53	8.6%	54	8.71%

2422

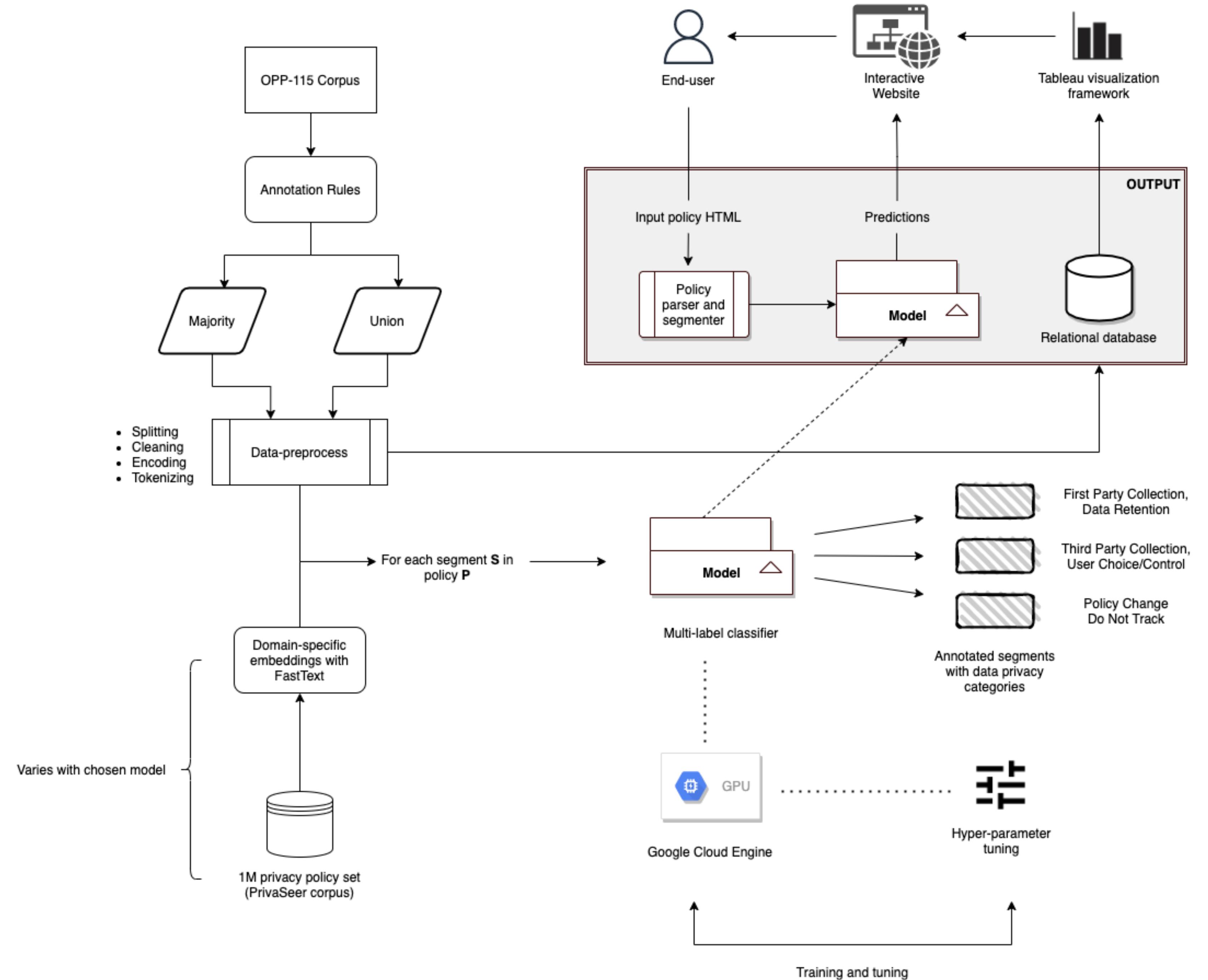
522

527

# INFORM

## Module

## Architecture



# Modeling Objective

- ★ Problem can be posed as a multi-label classification.
- ★ Produce probability  $P(i)$  for each category => predictions with an optimal threshold

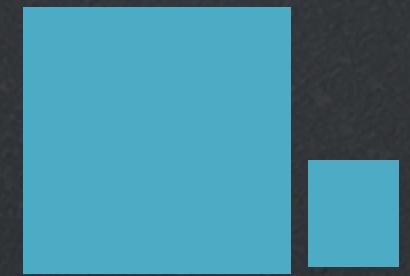
## Tradeoffs to consider



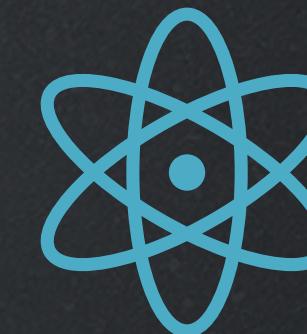
PERFORMANCE



LATENCY



SIZE



COMPUTE



INTERPRETABILITY

# Annotator Agreement

Category	Fleiss' Kappa
First Party Collection/Use	.76
Third Party Sharing/Collection	.76
Other	.49
User Choice/Control	.61
Data Security	.67
International and Specific Audiences	.87
User Access, Edit and Deletion	.74
Policy Change	.73
Data Retention	.55
Do Not Track	.91

Fleiss' Kappa statistic measures the degree of agreement between the annotators to label each segment with a privacy practice category

L

R

S

X

C

# Modeling Approach

- Primary focus, in this phase, is to establish baseline models
- Motivate the need to add complexity from data and model standpoints

## Advantages

- Low complexity => rapid experimentation via hyperparameter tuning
- Help discover data/code issues early on
- Leverage "the strength of weak learners"

# Modeling Baselines

## TRADITIONAL MACHINE LEARNING ALGORITHMS

Logistic Regression

Random Forest

LANGUAGE NEURAL MODEL

Convolutional Neural Network

XGBoost

Support Vector Machine (RBF)

# Traditional models

## Baselines

- ★ Built One-vs-Rest classifiers
- ★ Used bigram term frequency-inverse document frequency (tf-idf) encoding
- ★ Tuned with randomized search on the validation set



PERFORMANCE RESULTS



INTERPRETABLE RESULTS

# Performance Results

## Traditional Models

Logistic Regression

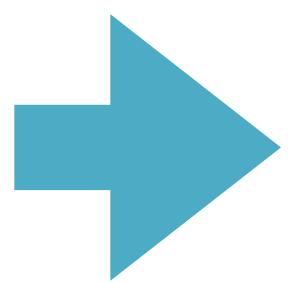
Random Forest

Support Vector Machine (RBF)

XGBoost



# Logistic Regression



## Performance Results

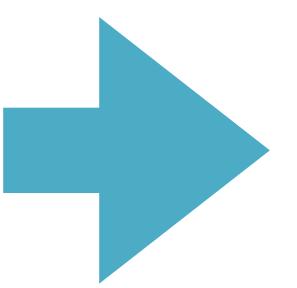
		precision	recall	f1-score	support	precision	recall	f1-score	support
International and Specific Audiences	Data Retention	0.52	0.57	0.54	23	0.31	0.45	0.37	11
	Data Security	0.71	0.79	0.75	56	0.79	0.71	0.75	31
	Do Not Track	0.71	0.83	0.77	6	0.44	1.00	0.62	4
	First Party Collection/Use	0.80	0.79	0.79	228	0.81	0.80	0.80	181
	Introductory/Generic	0.89	0.77	0.83	53	0.89	0.76	0.82	45
	Policy Change	0.64	0.57	0.60	122	0.68	0.67	0.68	58
	Practice not covered	0.70	0.72	0.71	29	0.65	0.83	0.73	18
	Privacy contact information	0.49	0.44	0.47	97	0.25	0.05	0.08	20
	Third Party Sharing/Collection	0.72	0.73	0.73	49	0.64	0.81	0.71	31
	User Access, Edit and Deletion	0.78	0.80	0.79	178	0.79	0.88	0.83	142
User Choice/Control	User Access, Edit and Deletion	0.76	0.65	0.70	34	0.77	0.74	0.76	23
	User Choice/Control	0.72	0.66	0.69	95	0.71	0.77	0.74	53
	micro avg	0.72	0.70	0.71	970	0.75	0.77	0.76	617
	macro avg	0.70	0.69	0.70	970	0.65	0.71	0.66	617
UNION	weighted avg	0.72	0.70	0.71	970	0.74	0.77	0.75	617
	samples avg	0.75	0.77	0.73	970	0.75	0.79	0.75	617

FK

UNION

MAJORITY

# Random Forest



## Performance Results

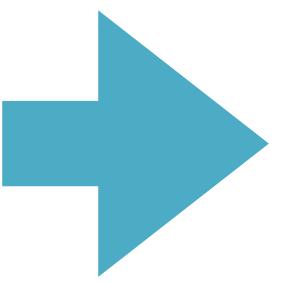
	precision	recall	f1-score	support	precision	recall	f1-score	support	
International and Specific Audiences	Data Retention	0.67	0.09	0.15	23	0.00	0.00	0.00	11
	Data Security	0.89	0.45	0.60	56	0.94	0.48	0.64	31
	Do Not Track	0.00	0.00	0.00	6	0.00	0.00	0.00	4
	First Party Collection/Use	0.88	0.68	0.77	228	0.90	0.71	0.79	181
	Introductory/Generic	0.94	0.62	0.75	53	0.96	0.51	0.67	45
	Policy Change	0.81	0.41	0.54	122	0.94	0.26	0.41	58
	Practice not covered	0.76	0.45	0.57	29	1.00	0.39	0.56	18
	Privacy contact information	0.80	0.12	0.21	97	0.00	0.00	0.00	20
	Third Party Sharing/Collection	1.00	0.47	0.64	49	0.81	0.42	0.55	31
	User Access, Edit and Deletion	0.89	0.71	0.79	178	0.90	0.68	0.78	142
User Choice/Control	User Access, Edit and Deletion	1.00	0.21	0.34	34	0.00	0.00	0.00	23
	User Choice/Control	0.76	0.29	0.42	95	0.85	0.43	0.58	53
	micro avg	0.87	0.49	0.63	970	0.90	0.52	0.66	617
	macro avg	0.78	0.38	0.48	970	0.61	0.32	0.41	617
Union	weighted avg	0.85	0.49	0.60	970	0.82	0.52	0.62	617
	samples avg	0.69	0.58	0.61	970	0.57	0.55	0.55	617

FK

UNION

MAJORITY

# Support Vector Machine (RBF)



## Performance Results

		precision	recall	f1-score	support	precision	recall	f1-score	support
International and Specific Audiences	Data Retention	0.56	0.39	0.46	23	0.60	0.27	0.37	11
	Data Security	0.78	0.71	0.75	56	0.81	0.68	0.74	31
	Do Not Track	1.00	0.67	0.80	6	1.00	0.50	0.67	4
	First Party Collection/Use	0.84	0.77	0.80	228	0.82	0.80	0.81	181
	Introductory/Generic	0.89	0.75	0.82	53	0.92	0.73	0.81	45
	Policy Change	0.63	0.55	0.59	122	0.77	0.64	0.70	58
	Practice not covered	0.73	0.76	0.75	29	0.84	0.89	0.86	18
	Privacy contact information	0.49	0.33	0.40	97	0.40	0.10	0.16	20
	Third Party Sharing/Collection	0.74	0.69	0.72	49	0.74	0.74	0.74	31
	User Access, Edit and Deletion	0.81	0.72	0.76	178	0.80	0.84	0.82	142
User Access, Edit and Deletion	User Choice/Control	0.81	0.50	0.62	34	1.00	0.70	0.82	23
	micro avg	0.70	0.55	0.62	95	0.78	0.74	0.76	53
	macro avg	0.75	0.64	0.69	970	0.81	0.74	0.77	617
	weighted avg	0.75	0.62	0.67	970	0.79	0.63	0.69	617
	samples avg	0.74	0.64	0.69	970	0.80	0.74	0.76	617

FK

UNION

MAJORITY



# XGBoost

## Performance Results

		precision	recall	f1-score	support	precision	recall	f1-score	support
International and Specific Audiences	Data Retention	0.59	0.43	0.50	23	0.50	0.18	0.27	11
	Data Security	0.80	0.62	0.70	56	0.74	0.65	0.69	31
	Do Not Track	0.80	0.67	0.73	6	0.50	1.00	0.67	4
	First Party Collection/Use	0.82	0.75	0.78	228	0.83	0.77	0.80	181
	Introductory/Generic	0.97	0.72	0.83	53	0.96	0.60	0.74	45
	Policy Change	0.70	0.50	0.58	122	0.87	0.57	0.69	58
	Practice not covered	0.74	0.69	0.71	29	0.78	0.78	0.78	18
	Privacy contact information	0.57	0.29	0.38	97	0.25	0.05	0.08	20
	Third Party Sharing/Collection	0.85	0.71	0.78	49	0.69	0.65	0.67	31
	User Access, Edit and Deletion	0.80	0.75	0.77	178	0.81	0.76	0.78	142
User Choice/Control	User Access, Edit and Deletion	0.83	0.44	0.58	34	0.86	0.52	0.65	23
	User Choice/Control	0.80	0.52	0.63	95	0.81	0.66	0.73	53
	micro avg	0.79	0.62	0.69	970	0.81	0.67	0.73	617
	macro avg	0.77	0.59	0.66	970	0.72	0.60	0.63	617
Union	weighted avg	0.78	0.62	0.68	970	0.80	0.67	0.72	617
	samples avg	0.75	0.70	0.69	970	0.70	0.69	0.69	617

FK

UNION

MAJORITY

# Interpretable Results

## Logistic Regression

y=First Party Collection/Use top features		y=Third Party Sharing/Collection top features		y=Privacy contact information top features		y=Policy Change top features		y=Do Not Track top features		y=User Choice/Control top features	
Weight?	Feature	Weight?	Feature	Weight?	Feature	Weight?	Feature	Weight?	Feature	Weight?	Feature
+16.744	collect	+27.070	share	+21.994	questions	+36.418	changes	+134.022	signals	+24.892	opt
+16.645	may use	+22.363	disclose	+19.908	please	+20.013	revised	+121.900	dnt	+22.342	unsubscribe
+12.229	uses	+20.309	companies	+19.740	contact	+19.492	updated	+84.742	track	+16.065	option
+12.070	use	+17.517	advertisers	+15.551	com	+17.474	october	+62.955	track	+15.043	consent
+11.795	participate	+16.633	shared	+13.437	eu	+17.474	last	+52.392	signals	+12.472	want
+11.512	emails	+16.535	sell	+11.662	contacting	+14.102	changes	+44.963	response	+12.298	disable
+11.253	commercial	+16.471	partners	+10.923	suite	+13.742	privacy	+44.385	signals	+11.733	choices
... 5476 more positive ...		+14.795	partner	+10.737	please send	+12.870	notice	+43.983	meaning	+11.608	written
... 4531 more negative ...		+12.951	providers may	+10.601	contacting us	+12.295	modified june	+43.720	track dnt	+11.462	options
-11.174	please	... 5053 more positive ...		+10.454	choose opt	+12.164	last updated	+35.720	currently	... 5117 more positive ...	
-11.417	secure	... 3042 more negative ...		... 3141 more positive ...		+11.821	policy time	+35.720	assign	... 2334 more negative ...	
-14.348	privacy	-12.960	security	... 1153 more negative ...		... 1642 more positive ...		+35.720	meaning	-11.341	track
						... 415 more negative ...		... 154 more positive ...		... 703 more negative ...	

Top ranked vocabulary for each category

# Interpretable Results

## Logistic Regression

y=First Party Collection/Use (probability 0.921, score 2.463) top features

Contribution?	Feature
+4.876	Highlighted in text (sum)
-2.413	<BIAS>

collection personally identifiable information social media sites addition , interact atlantic property page account social media platform , facebook , twitter , tumblr , linkedin , may collect personally identifiable information make available us page account including social media account id . however , comply privacy policies corresponding social media platform collect store personally identifiable information permitted collect social media platforms . choose link login atlantic account social networking service , atlantic service may share certain information activities . consent , also may share information activities , including view sites , social networks users .

y=Third Party Sharing/Collection (probability 0.995, score 5.364) top features

Contribution?	Feature
+8.626	Highlighted in text (sum)
-3.262	<BIAS>

collection personally identifiable information social media sites addition , interact atlantic property page account social media platform , facebook , twitter , tumblr , linkedin , may collect personally identifiable information make available us page account including social media account id . however , comply privacy policies corresponding social media platform collect store personally identifiable information permitted collect social media platforms . choose link login atlantic account social networking service , atlantic service may share certain information activities . consent , also may share information activities , including view sites , social networks users .

Token importance attributions for First Party/ Third Party categories

# Interpretable Results

## Logistic Regression

y=Privacy contact information (probability 0.001, score -7.457) top features

Contribution?	Feature
-2.549	Highlighted in text (sum)
-4.908	<BIAS>

collection personally identifiable information social media sites addition , interact atlantic property page account social media platform , facebook , twitter , tumblr , linkedin , may collect personally identifiable information make available us page account including social media account id . however , comply privacy policies corresponding social media platform collect store personally identifiable information permitted collect social media platforms . choose link login atlantic account social networking service , atlantic service may share certain information activities . consent , also may share information activities , including view sites , social networks users .

y=Policy Change (probability 0.000, score -7.690) top features

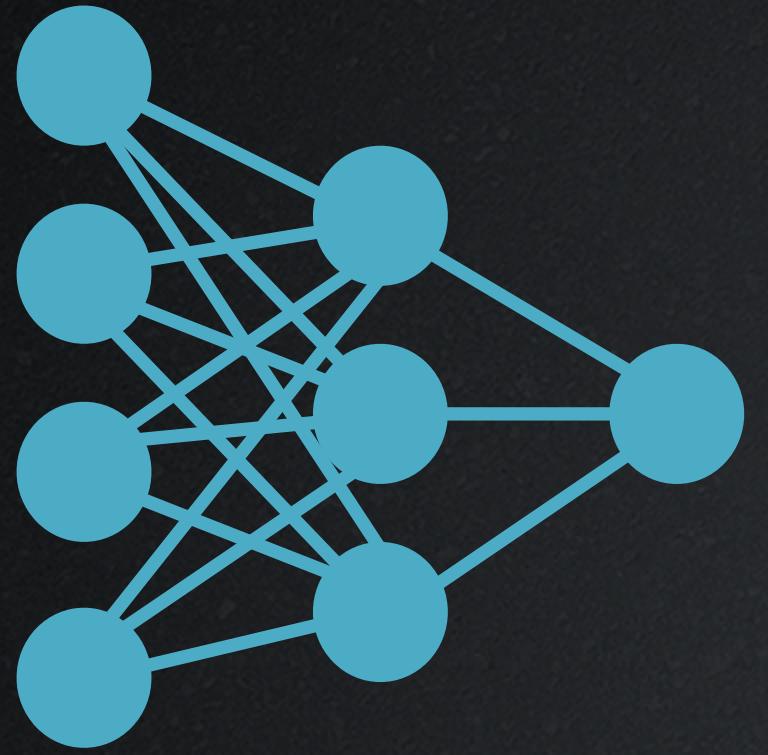
Contribution?	Feature
-1.229	Highlighted in text (sum)
-6.461	<BIAS>

collection personally identifiable information social media sites addition , interact atlantic property page account social media platform , facebook , twitter , tumblr , linkedin , may collect personally identifiable information make available us page account including social media account id . however , comply privacy policies corresponding social media platform collect store personally identifiable information permitted collect social media platforms . choose link login atlantic account social networking service , atlantic service may share certain information activities . consent , also may share information activities , including view sites , social networks users .

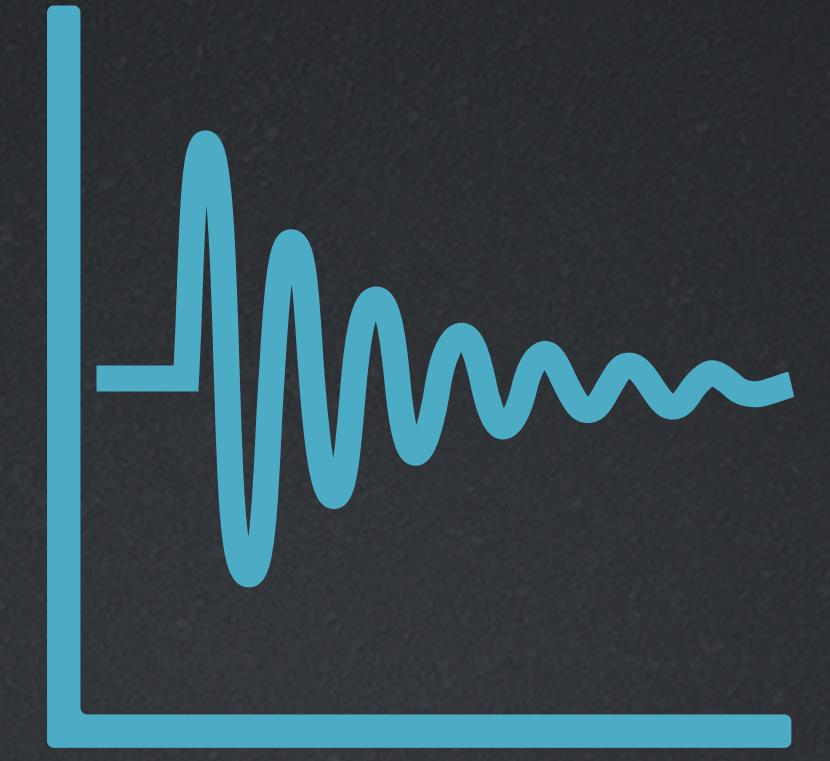
Token importance attributions for Privacy Contact/ Policy Change categories

# Language Neural Model

## Convolutional Neural Network



ARCHITECTURE



EXPERIMENTS

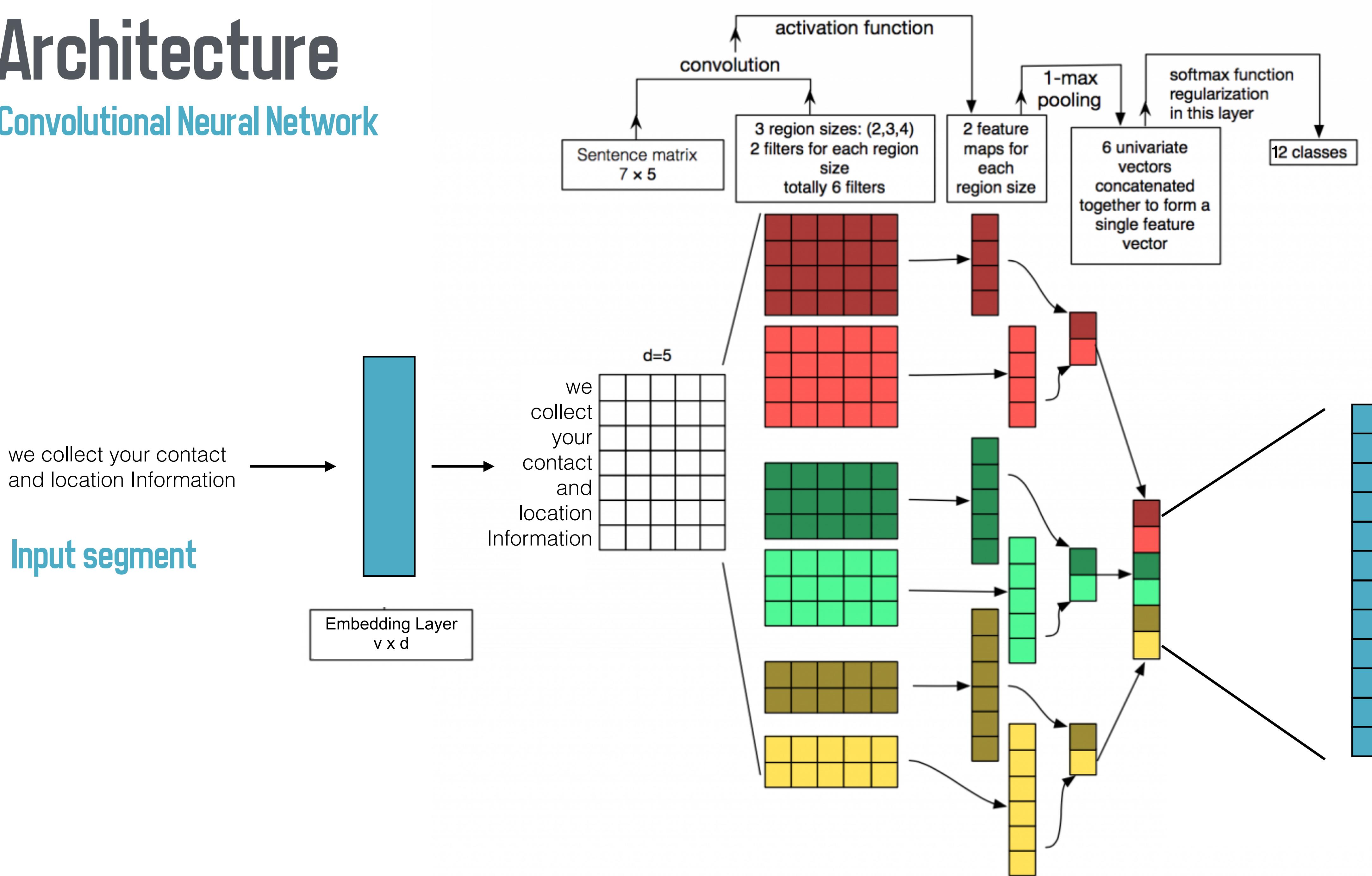


PERFORMANCE



# Architecture

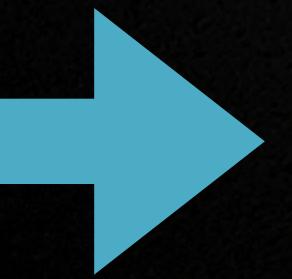
## Convolutional Neural Network



Source

# Experiments

## Convolutional Neural Network



## Hyperparameters

Token level

Embedding Type

Learning Rate

Number of Kernels

Lower Text

Trainable Embedding

Hidden Dimensions

Stem Text

Embedding Dimension

Number of  
Epochs

Filter Size

Dropout

DATA

EMBEDDING

TRAINING

ARCHITECTURE

# Ablation Study

## Convolutional Neural Network

MLFlow

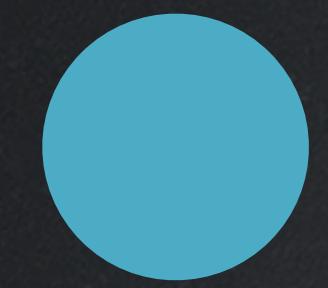


- Optuna library for randomized search
- MLFlow library for experiment tracking
- GCP instance (Nvidia Tesla T4)
- CUDA and CUDNN libraries

CNN Ablation									
Experiment ID	CHAR LEVEL	Embeddings	Trainable	Embed Dim	Union dataset	Tuning params	Complete		
CNN_C_FE_U	<input checked="" type="checkbox"/>	Random	True	Tuned	<input checked="" type="checkbox"/>	MFS ED NF HD DP LR EP	<input checked="" type="checkbox"/>		
CNN_C_FE_M	<input checked="" type="checkbox"/>	Random	True	Tuned	<input type="checkbox"/>	MFS ED NF HD DP LR EP	<input checked="" type="checkbox"/>		
CNN_W_FE_U	<input type="checkbox"/>	Random	True	Tuned	<input checked="" type="checkbox"/>	MFS ED NF HD DP LR EP	<input checked="" type="checkbox"/>		
CNN_W_FE_M	<input type="checkbox"/>	Random	True	Tuned	<input type="checkbox"/>	MFS ED NF HD DP LR EP	<input checked="" type="checkbox"/>		
CNN_W_PEG_U_FR	<input type="checkbox"/>	Pre-trained (GloVe)	False	300	<input checked="" type="checkbox"/>	MFS NF HD DP LR EP	<input checked="" type="checkbox"/>		
CNN_W_PEG_U_TR	<input type="checkbox"/>	Pre-trained (GloVe)	True	300	<input checked="" type="checkbox"/>	MFS NF HD DP LR EP	<input checked="" type="checkbox"/>		
CNN_W_PEG_M_FR	<input type="checkbox"/>	Pre-trained (GloVe)	False	300	<input type="checkbox"/>	MFS NF HD DP LR EP	<input type="checkbox"/>		
CNN_W_PEF_U_TR	<input type="checkbox"/>	Pre-trained (fastText)	True	300	<input checked="" type="checkbox"/>	MFS NF HD DP LR EP	<input type="checkbox"/>		
CNN_W_PEF_U_FR	<input type="checkbox"/>	Pre-trained (fastText)	False	300	<input checked="" type="checkbox"/>	MFS NF HD DP LR EP	<input type="checkbox"/>		
CNN_W_PEF_M	<input type="checkbox"/>	Pre-trained (fastText)			<input type="checkbox"/>	MFS NF HD DP LR EP	<input type="checkbox"/>		
CNN_W_DEF_U	<input type="checkbox"/>	Domain-specific			<input checked="" type="checkbox"/>	MFS NF HD DP LR EP	<input type="checkbox"/>		
CNN_W_DEF_M	<input type="checkbox"/>	Domain-specific			<input type="checkbox"/>	MFS NF HD DP LR EP	<input type="checkbox"/>		

# Performance Results

## Convolutional Neural Network



CHARACTER LEVEL

Used upto 10 character n-grams



WORD LEVEL

Used upto 5 word n-grams

# Performance Results

FK

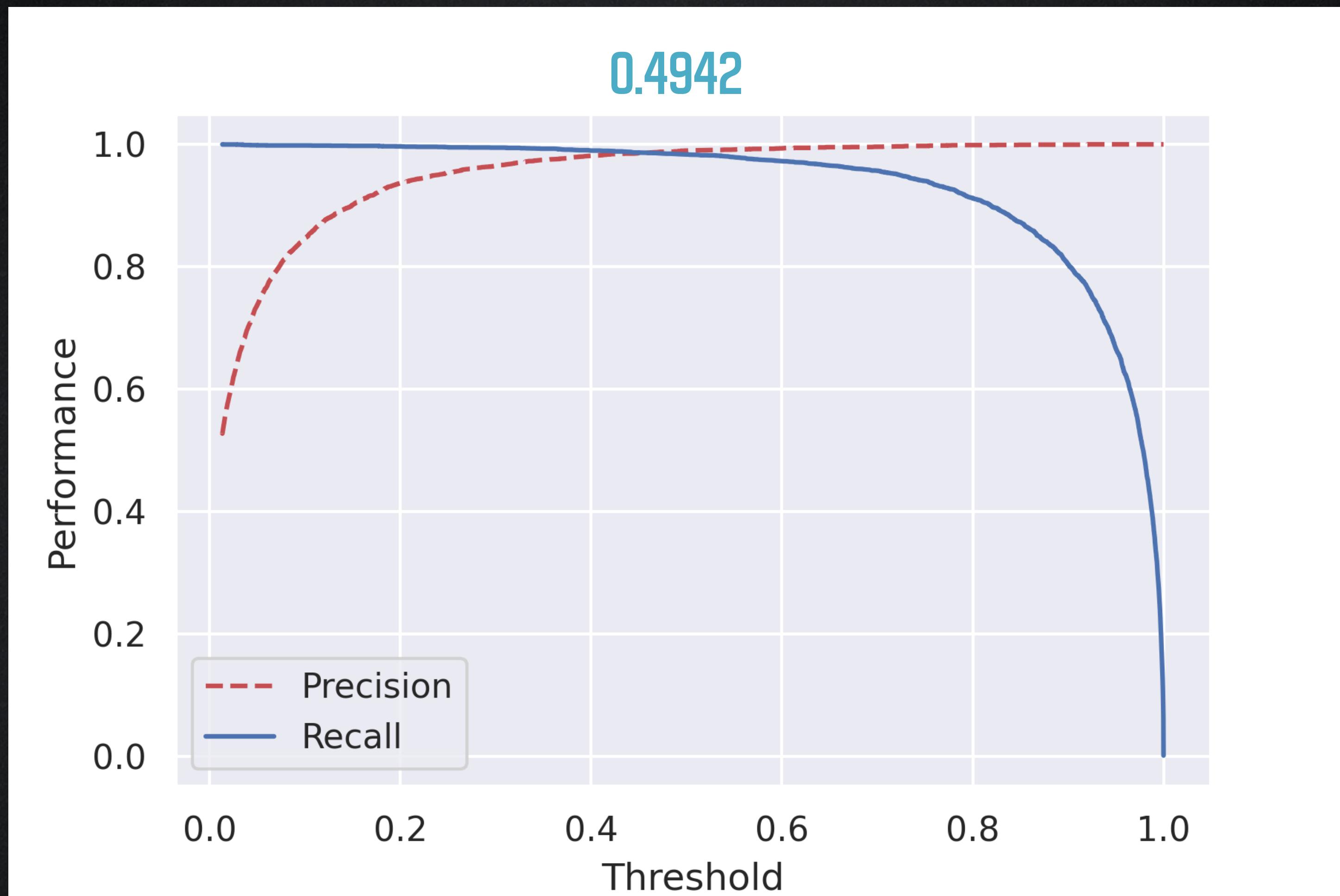
Character level → ← Word level

← Union → ← Majority → ← Union → ← Majority →

	precision	recall	f1-score	support												
<b>Data Retention</b>	0.800000	0.347826	0.484848	23.0	1.000000	0.083333	0.153846	12.0	1.000000	0.173913	0.296296	23.0	1.000000	0.083333	0.153846	12.0
<b>Data Security</b>	0.730769	0.666667	0.697248	57.0	0.923077	0.774194	0.842105	31.0	0.853659	0.614035	0.714286	57.0	0.956522	0.709677	0.814815	31.0
<b>Do Not Track</b>	1.000000	1.000000	1.000000	5.0	1.000000	0.800000	0.888889	5.0	1.000000	0.200000	0.333333	5.0	1.000000	0.800000	0.888889	5.0
<b>First Party Collection/Use</b>	0.710714	0.872807	0.783465	228.0	0.783505	0.839779	0.810667	181.0	0.835681	0.780702	0.807256	228.0	0.804233	0.839779	0.821622	181.0
<b>International and Specific Audiences</b>	0.960000	0.905660	0.932039	53.0	0.900000	0.800000	0.847059	45.0	1.000000	0.849057	0.918367	53.0	0.970588	0.733333	0.835443	45.0
<b>Introductory/Generic</b>	0.636364	0.573770	0.603448	122.0	0.767442	0.559322	0.647059	59.0	0.764706	0.532787	0.628019	122.0	0.833333	0.508475	0.631579	59.0
<b>Policy Change</b>	0.850000	0.586207	0.693878	29.0	0.937500	0.833333	0.882353	18.0	0.933333	0.482759	0.636364	29.0	0.833333	0.833333	0.833333	18.0
<b>Practice not covered</b>	0.515152	0.350515	0.417178	97.0	0.000000	0.000000	0.000000	20.0	0.606061	0.206186	0.307692	97.0	1.000000	0.050000	0.095238	20.0
<b>Privacy contact information</b>	0.833333	0.625000	0.714286	48.0	0.791667	0.633333	0.703704	30.0	0.911765	0.645833	0.756098	48.0	0.800000	0.533333	0.640000	30.0
<b>Third Party Sharing/Collection</b>	0.683486	0.837079	0.752525	178.0	0.698718	0.767606	0.731544	142.0	0.829114	0.735955	0.779762	178.0	0.787879	0.732394	0.759124	142.0
<b>User Access, Edit and Deletion</b>	0.857143	0.529412	0.654545	34.0	0.882353	0.652174	0.750000	23.0	0.937500	0.441176	0.600000	34.0	0.875000	0.304348	0.451613	23.0
<b>User Choice/Control</b>	0.686275	0.736842	0.710660	95.0	0.809524	0.629630	0.708333	54.0	0.698795	0.610526	0.651685	95.0	0.861111	0.574074	0.688889	54.0
<b>micro avg</b>	0.707216	0.707946	0.707581	969.0	0.783688	0.712903	0.746622	620.0	0.820055	0.616099	0.703595	969.0	0.828685	0.670968	0.741533	620.0
<b>macro avg</b>	0.771936	0.669315	0.703677	969.0	0.791149	0.614392	0.663797	620.0	0.864218	0.522744	0.619097	969.0	0.893500	0.558507	0.634533	620.0
<b>weighted avg</b>	0.708186	0.707946	0.697765	969.0	0.769452	0.712903	0.728699	620.0	0.814190	0.616099	0.685493	969.0	0.842847	0.670968	0.720337	620.0
<b>samples avg</b>	0.743398	0.770745	0.720586	969.0	0.721063	0.737192	0.712081	620.0	0.799002	0.701585	0.713037	969.0	0.719165	0.695762	0.693169	620.0

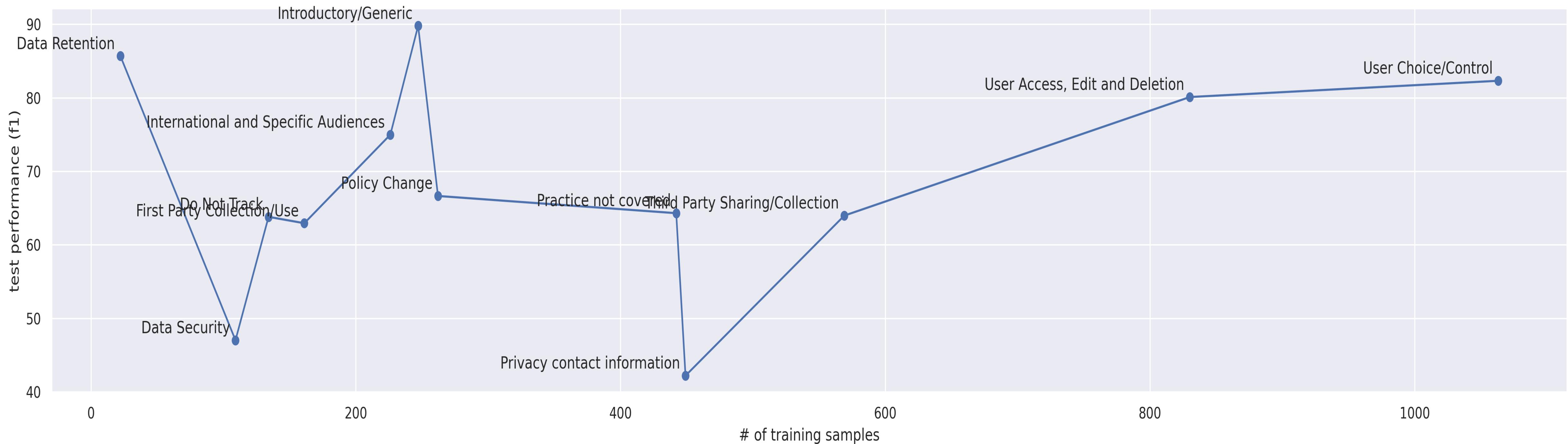
# Performance Results

## Finding Optimal Threshold



# Performance Results

## Number of samples vs F1



# Performance Results (CNN - word and union)

Taking a closer look at the predictions for each category

First Party Collection/Use

Third Party Sharing/  
Collection

User Choice/Control

Data Security

Data Retention

Do Not Track

Policy Change

User Access/Edit/Deletion

International and Specific  
Audiences

Practice Not Covered

Privacy Contact Information

Introductory/Generic

# First Party Collection/Use

Precision	Recall	Weighted F1	Support
80.93	83.77	82.32	228

True Positives

False Positives

False Negatives

# True Positives

## First Party Collection/Use

==== True positives ===

information that sci news com may collect online sci news com may collect and process the following data about you information that you provide by filling in forms on our site including names e mail and website addresses we may also ask you for information for other purposes for example when you report a problem with our site

true: ['First Party Collection/Use']

pred: ['First Party Collection/Use']

information collection and use information collection redorbit inc is the sole owner of the information collected on redorbit com redorbit inc collects information from our users at several different points on our web site

true: ['First Party Collection/Use', 'Introductory/Generic', 'Practice not covered']

pred: ['First Party Collection/Use', 'Introductory/Generic']

log files like most standard web site servers we use log files this includes internet protocol ip addresses browser type internet service provider isp referringexit pages platform type date/time stamp and number of clicks to analyze trends administer the site track user's movement in the aggregate and gather broad demographic information for aggregate use ip addresses etc are not linked to personally identifiable information we use a tracking utility called clicktracks that uses log files to analyze user movement

true: ['First Party Collection/Use']

pred: ['First Party Collection/Use']

# False Positives

## First Party Collection/Use

==== False positives ===

aggregate information non personally identifiable we share aggregated demographic information with our partners and advertisers this is not linked to any personally identifiable information

true: ['Third Party Sharing/Collection']

pred: ['First Party Collection/Use', 'Third Party Sharing/Collection']

some of our services aol search and netscape isp for example may also offer you the ability to manage and control information collected or used when you use these services

true: ['User Choice/Control']

pred: ['First Party Collection/Use']

specific web site provisions this section below addresses this sites unique functionality and offerings specific to particular areas of each site and serves as a supplement to the general provisions above

true: ['Introductory/Generic']

pred: ['First Party Collection/Use']

# False Negatives

## First Party Collection/Use

==== False negatives ====

newsletter if a user wishes to subscribe to our newsletter we ask for contact information such as name and email address out of respect for our users privacy we provide a way to opt out of these communications please see the choice and opt out sections

true: ['First Party Collection/Use']

pred: ['User Choice/Control']

newsletter if a user wishes to subscribe to our newsletter we ask for contact information such as name and email address out of respect for our users privacy we provide a way to opt out of these communications please see the choice and opt out sections

true: ['First Party Collection/Use']

pred: ['User Choice/Control']

other information please note that aol may use information about your use of certain aol communication tools for example aol mail and aol instant messenger however when you use aol communication tools aol does not read your private online communications without your consent

true: ['First Party Collection/Use']

pred: ['User Choice/Control']

# Future Directions



Advanced interpretation approaches



Scale up model and data complexities



QUERY Module



Implementation of web framework



# Scale up model and data complexities

## Future Directions



Pre-trained general purpose embeddings (GloVe, fastText)



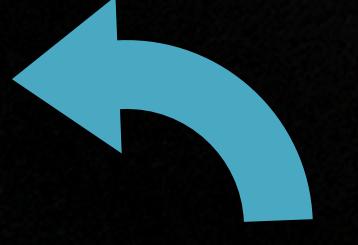
Domain specific embeddings using 1M privacy policies (PrivSeer Corpus)



Sequence Models

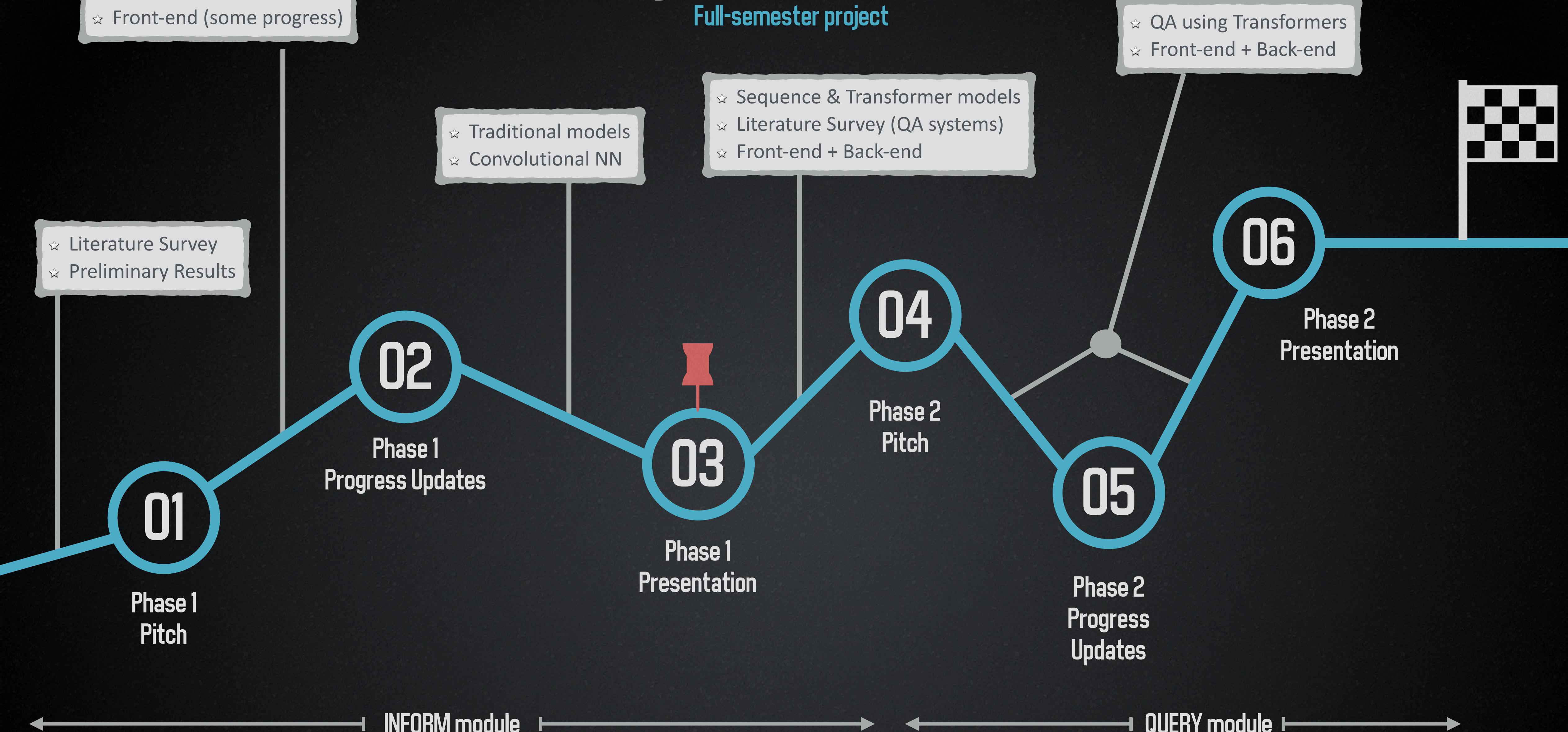


Transformers (BERT, LEGAL-BERT, fine-tuned BERT)



# Project Timeline

Full-semester project



THANK  
YOU

